

VU Research Portal

Optimal measurement conditions for EEG/MEG source analysis

Huizenga, H.M.; Heslenfeld, D.J.; Molenaar, P.C.M.

published in
Psychometrika
2002

DOI (link to publisher)
[10.1007/BF02294848](https://doi.org/10.1007/BF02294848)

document version
Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Huizenga, H. M., Heslenfeld, D. J., & Molenaar, P. C. M. (2002). Optimal measurement conditions for EEG/MEG source analysis. *Psychometrika*, 67(2), 299-313. <https://doi.org/10.1007/BF02294848>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

OPTIMAL MEASUREMENT CONDITIONS FOR SPATIOTEMPORAL EEG/MEG SOURCE ANALYSIS

HILDE M. HUIZENGA

UNIVERSITY OF AMSTERDAM

DIRK J. HESLENFELD

UNIVERSITY OF AMSTERDAM AND
FREE UNIVERSITY OF AMSTERDAM

PETER C.M. MOLENAAR

UNIVERSITY OF AMSTERDAM

Electromagnetic source analysis yields estimates of the sources of the Electro- and/or MagnetoEncephaloGram (EEG/MEG) and thus generates a functional description of the human brain. The standard errors of the source estimates are influenced by the number and position of EEG/MEG sensors, by the number of time samples, and by the number of trials in which EEG/MEG is measured. Therefore, optimal design theory is applied to determine the required number and position of sensors, the required number of samples, and the required number of trials. To that end, the Fedorov exchange algorithm is extended to incorporate multi-response models. A simulation study and an empirical study on visual evoked potentials indicate that the proposed method is fast and reliable, and improves source precision considerably.

Key words: spatiotemporal EEG/MEG source analysis, multiresponse nonlinear regression, optimal design, Fedorov exchange

Introduction

Neuroscientific research has benefited greatly from the development of noninvasive functional imaging techniques like Positron Emission Tomography (e.g., Vardi, Shepp & Kaufman, 1985), functional Magnetic Resonance Imaging (e.g., Lange & Zeger, 1997), and ElectroMagnetic Source Analysis (e.g., Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). Source analysis is founded on the fact that active brain areas generate electromagnetic activity, which can be measured with the Electro- and/or Magneto-EncephaloGram (EEG/MEG). A source analysis offers the opportunity to estimate the sources of this EEG/MEG, and thus to identify active brain areas.

Electromagnetic source analysis is in general based on nonlinear regression analysis. The regression model defines the relation between the dependent variable (EEG/MEG), the independent variable (EEG/MEG sensor characteristics), and the parameters defining the sources. The standard errors of the source parameter estimates are affected by the number of trials in which EEG/MEG is measured, by the number of time samples, and by the number and position of EEG/MEG sensors. Therefore we apply optimal design theory (e.g., Atkinson & Donev, 1992) to determine the required number of trials, the required number of time samples, and the required number and position of sensors.

There are several reasons why such a method is needed. First, source analysis is often performed on averages of many trials to decrease the noise level of the data, and thus the standard

We would like to thank Juha Virtanen for his help in collecting the empirical data. The research of H.M.H. has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. D.J.H. was supported by grant number 575-65-058 from the Dutch Organization for Scientific Research (NWO).

Requests for reprints should be sent to Hilde Huizenga, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018WB Amsterdam, THE NETHERLANDS. E-Mail: op.huizenga@macmail.psy.uva.nl

error of the parameter estimates. However, the number of trials should not be too high, since this may introduce learning or habituation effects. Therefore, it is useful to determine the minimal number of required trials. Second, EEG/MEG is measured in continuous time, but online sampled in discrete samples. Only a limited time interval, expected to be related to the process of interest, is then analyzed. Increasing the number of samples, either by increasing the sampling density or by increasing the time interval, will in general yield more precise source estimates. However, a too high sampling density may generate data storage difficulties. In addition, a long time interval may require more sources to be estimated, and thus the estimates may even become less precise. For these reasons, it is useful to determine the required number of samples. Third, although there do exist devices with up to 250 sensors, in most research only a limited number of sensors is used. Either there is only a limited number of sensors available, which are then placed on areas expected to be most informative (e.g., Gunji, Kagigi, & Hoshiyama, 2000; Kenemans, Baas, Mangum, Lijffijt, & Verbaten, 2000). Or many sensors are measured, but only a subset, expected to be most informative, is used for source analysis (e.g., Tiitinen, Sivonen, Alku, Virtanen, & Näätänen, 1999). Therefore, it is useful to determine the minimal number of sensors required, and their optimal positions.

There exist two approaches to determine the required number, and in some instances, position of sensors. In the Monte Carlo approach the best sensor configuration is chosen by comparing simulation results on several configurations (Cuffin, 1985; Gaumond, Lin, & Geselowitz, 1983; Hari, Joutsiniemi, & Sarvas, 1988; Huizenga & Molenaar, 1994; Ogura & Sekihara, 1993). This is very time consuming, and therefore not very useful in practice. The second approach yields an estimate of the maximally allowed distance between sensors to prevent aliasing of high spatial frequencies (Ahonen, Hämäläinen, Ilmoniemi, Kajola, Knuutila, Simola, & Vilkmann, 1993; Hämäläinen et al., 1993; Kuc, 1996; Nunez, 1988; Spitzer, Cohen, Fabrikant, & Hallet, 1989; Vaidyanathan & Buckley, 1997). This is the method of choice if many sensors can be measured, but it is less useful if the number of sensors is limited. The present method supplements these traditional approaches. The method is analytic, and therefore fast. Moreover, it is specifically suited for situations where only a limited number of sensors is measured and/or analyzed.

The paper is organized as follows. First, we give a brief introduction to source analysis. Then we describe the optimal design technique, and we report a simulation study on its validity and merits. Finally, we report an empirical illustration in a study on visual evoked potentials.

Electromagnetic Source Analysis: Physical Aspects

The following conventions are used: scalars are *italics*, matrices are in **bold uppercase**, and vectors are **columns** and in **boldface**. The prime symbol (i.e., $'$) denotes transposition, a superscript of minus one (i.e., $^{-1}$) denotes inversion, " $\|\cdot\|$ " denotes vector norm, " $\text{vec}(\cdot)$ " denotes the vector operator, " $|\cdot|$ " denotes the determinant, and " \otimes " denotes the Kronecker product.

EEG and MEG measure respectively the electric potential and the magnetic field generated by active brain areas. Given these measurements, and given a suitable model for these measurements, it is possible to estimate the characteristics of active brain areas, most importantly their location and amplitude. A suitable model of the measurements should (a) give an adequate description of active brain areas, and (b) give an adequate description of the conductive properties of the head. In the following we describe these two aspects briefly.

Active brain areas can be modeled by current dipoles. A current dipole consists of current flowing between poles that are separated an infinitesimal small distance. Such a dipole can be characterized by its location $\boldsymbol{\tau} = [\tau_1, \tau_2, \tau_3]'$ and moment $\boldsymbol{\xi} = [\xi_1, \xi_2, \xi_3]'$. The location vector consists of Euclidean x , y and z coordinates. Positive x points towards theinion (in the neck), positive y points towards the right ear, and positive z points upwards. The moment vector is described in a parallel system, but with zero point at source location. This parameterization will be called the Euclidean parametrization. An alternative description for the moment vector, which

will be called the spherical parametrization, consists of two angles that describe source orientation, and a parameter ψ that describes source amplitude. A third way to describe a dipole is the normalized Euclidean parametrization. Let the source location vector be $\boldsymbol{\tau} = \varphi \boldsymbol{\tau}^n$, where $\boldsymbol{\tau}^n = [\tau_1^n, \tau_2^n, \tau_3^n]'$ denote the 3 location direction cosines and where φ denotes source eccentricity. Let the source moment vector be $\boldsymbol{\xi} = \psi \boldsymbol{\xi}^n$, where $\boldsymbol{\xi}^n = [\xi_1^n, \xi_2^n, \xi_3^n]'$ denote the 3 orientation direction cosines and where ψ denotes source amplitude (e.g., Huizenga & Molenaar, 1994). These parameters should satisfy the two nonlinear constraints $\tau_1^{n2} + \tau_2^{n2} + \tau_3^{n2} = 1$ and $\xi_1^{n2} + \xi_2^{n2} + \xi_3^{n2} = 1$.

A current dipole generates an electric potential and a magnetic field, which are distorted by brain tissues. For example the low skull conductivity smears and attenuates the electric potential, and, to a lesser extent, the magnetic field. Therefore it is necessary to model the conductive properties of the head. Head models range from the realistically shaped model which is computationally very demanding, to the homogenous sphere which is computationally simple. Most often a compromise is sought in the model with four concentric spheres. These spheres model the different conductivities of the brain, cerebrospinal fluid, skull and scalp. A three-sphere model in which the fluid is omitted is also quite common.

In this paper we focus on EEG, that is, on measurements of the electric potential. The electric potential at sensor n is defined as $y_n = y_n^* - y_r^*$, where y_n^* is activity at sensor n , and y_r^* is reference activity. This reference activity can be measured on any scalp position, often it is measured near one of the ears. A reference $y_r^* = 0$ will be called a zero reference.

The nonlinear function describing the electric potential generated by one dipole in four concentric isotropic spheres was given by Cuffin and Cohen (1979) and Mosher, Spencer, Leahy, and Lewis (1993). The potential generated by multiple dipoles is simply the sum of potentials generated by separate dipoles. Let the radii of the inner to outer spheres be r_1, r_2, r_3, r_4 , and let the matching conductivities be c_1, c_2, c_3, c_4 . The electric potential generated on a sensor \mathbf{x} measured with respect to a reference sensor \mathbf{x}_r is

$$f(\mathbf{x}, \mathbf{x}_r, \boldsymbol{\tau}, \boldsymbol{\xi}) = \boldsymbol{\xi}' \mathbf{d}(\mathbf{x}, \boldsymbol{\tau}) - \boldsymbol{\xi}' \mathbf{d}(\mathbf{x}_r, \boldsymbol{\tau}) \quad (1)$$

with

$$\mathbf{d}(\mathbf{x}, \boldsymbol{\tau}) = \sum_{j=1}^{\infty} w_j \left(\frac{\|\boldsymbol{\tau}\|}{r_4} \right)^{j-1} (\mathbf{a}_1 P_j^1 \cos \alpha \cos \beta + \mathbf{a}_2 P_j^1 \cos \alpha \sin \beta + \mathbf{a}_3 j P_j \cos \alpha).$$

The vectors $\mathbf{a}_1, \mathbf{a}_2$ and \mathbf{a}_3 are the basis for rotated coordinate axes that place the dipole on the z -axis, that is $\mathbf{a}_1' \boldsymbol{\tau} = 0, \mathbf{a}_2' \boldsymbol{\tau} = 0, \mathbf{a}_3' \boldsymbol{\tau} = \|\boldsymbol{\tau}\|$. The angles α and β are the longitude and latitude of the sensor in the rotated coordinate system. P_j is the j -th order Legendre polynomial, and P_j^1 is the j -th order associated Legendre polynomial. Let $\tilde{r}_1 = r_1/r_4; \tilde{r}_2 = r_2/r_4; \tilde{r}_3 = r_3/r_4$ and $\tilde{c}_1 = c_1/c_2; \tilde{c}_2 = c_2/c_3; \tilde{c}_3 = c_3/c_4$ then

$$w_j = \frac{1}{4\pi c_4 r_4^2} \frac{(2j+1)^4 (\tilde{r}_2 \tilde{r}_3)^{2j+1}}{j \Gamma_j} \quad \text{where } \Gamma_j \text{ equals} \quad (2)$$

$$\begin{aligned} & \tilde{r}_3^{2j+1} \left\{ \tilde{r}_1^{2j+1} j (\tilde{c}_1 - 1) (\tilde{c}_2 - 1) (j+1) + \tilde{r}_2^{2j+1} (\tilde{c}_1 j + j+1) (\tilde{c}_2 j + j+1) \right\} \\ & \left\{ (\tilde{c}_3 j + j+1) + (j+1) (\tilde{c}_3 - 1) \tilde{r}_3^{2j+1} \right\} \\ & + (j+1) \tilde{r}_2^{2j+1} \left\{ \tilde{r}_1^{2j+1} (\tilde{c}_1 - 1) (\tilde{c}_2 j + \tilde{c}_2 + j) + \tilde{r}_2^{2j+1} (\tilde{c}_1 j + j+1) (\tilde{c}_2 - 1) \right\} \\ & \left\{ j (\tilde{c}_3 - 1) + (\tilde{c}_3 j + \tilde{c}_3 + j) \tilde{r}_3^{2j+1} \right\}. \end{aligned}$$

Electromagnetic Source Analysis: Statistical Aspects

In this section we outline how source parameters can be estimated from EEG/MEG data. These data consist of measurements on T time samples, N sensors and M trials. The data of M trials are averaged to improve the signal to noise ratio. Two general approaches can be adopted to estimate the sources. In an instantaneous analysis, each sample is analyzed separately. In a spatiotemporal analysis, T samples are analyzed simultaneously under the assumption that some source parameters are fixed over samples. In the following we describe a spatiotemporal analysis, the instantaneous case can be obtained by setting $T = 1$.

Spatiotemporal analysis is based on the multi (T) response nonlinear regression model $\mathbf{Y} = \mathbf{F}(\mathbf{X}, \boldsymbol{\theta}) + \mathbf{E}$, where \mathbf{Y} , \mathbf{F} and \mathbf{E} are N by T matrices with measurements, model and noise respectively. $\mathbf{F}(\mathbf{X}, \boldsymbol{\theta})$ is the model, for example equation (1), \mathbf{X} is the matrix with sensor coordinates, and $\boldsymbol{\theta}$ is the vector with source parameters. It is assumed that source location and orientation are fixed in time, whereas their amplitude may vary. Since source amplitude appears linearly (see (1)), \mathbf{F} can be written as $\mathbf{H}\boldsymbol{\Psi}$. \mathbf{H} is the N by D "lead field" matrix with signals generated by D sources with unit amplitude. $\boldsymbol{\Psi}$ is the D by T matrix with time varying amplitude of D sources. In the spherical parametrization $\boldsymbol{\theta}$ thus contains $P = (3 + 2 + T)D$ source parameters: for each source 3 parameters to define location, 2 parameters to define orientation, and T parameters to define time varying source amplitude.

In general it is assumed that noise is homoscedastic and uncorrelated. The parameters $\boldsymbol{\theta}$ are then estimated by minimizing the ordinary least squares (OLS) function $s(\boldsymbol{\theta}) = [\text{vec}(\mathbf{F}(\mathbf{X}, \boldsymbol{\theta})) - \text{vec}(\mathbf{Y})]'[\text{vec}(\mathbf{F}(\mathbf{X}, \boldsymbol{\theta})) - \text{vec}(\mathbf{Y})]$. Let $s(\hat{\boldsymbol{\theta}})$ denote its minimum, then the estimated noise variance is $\hat{\sigma}^2 = s(\hat{\boldsymbol{\theta}})/(NT - P)$. The covariance matrix of the estimated parameters $\hat{\boldsymbol{\theta}}$ is (see Huizenga & Molenaar, 1994)

$$\hat{\mathbf{C}} = \hat{\sigma}^2[\hat{\mathbf{G}}'\hat{\mathbf{G}}]^{-1}. \quad (3)$$

$\hat{\mathbf{G}}$ is the NT by P matrix with first order partial derivatives of $\text{vec}(\mathbf{F}(\mathbf{X}, \boldsymbol{\theta}))$ with respect to the P parameters. Equation (3) is a matrix with P nonzero positive eigenvalues.

Tests on the source parameters can be constructed in the following way. Let the vector $\mathbf{r}(\hat{\boldsymbol{\theta}})$ contain Q (non)linear hypotheses on $\hat{\boldsymbol{\theta}}$ and let the Q by P matrix $\mathbf{R}(\hat{\boldsymbol{\theta}})$ contain first order partial derivatives of Q hypotheses with respect to P parameters. Then the test statistic for $\mathbf{r}(\hat{\boldsymbol{\theta}}) - \mathbf{q} = \mathbf{0}$ is (Seber & Wild, 1989, p. 198, 228, 703)

$$[\mathbf{r}(\hat{\boldsymbol{\theta}}) - \mathbf{q}]'[\mathbf{R}(\hat{\boldsymbol{\theta}})\hat{\mathbf{C}}\mathbf{R}(\hat{\boldsymbol{\theta}})']^{-1}[\mathbf{r}(\hat{\boldsymbol{\theta}}) - \mathbf{q}] \sim QF_{Q;NT-P}^{\alpha}, \quad (4)$$

where F denotes the $1 - \alpha$ quantile of the F distribution with Q and $NT - P$ degrees of freedom.

In case of the normalized Euclidean parametrization, the parameter vector $\boldsymbol{\theta}^c$ contains $P^c = (4 + 3 + T)D$ parameters, which should satisfy the aforementioned $2D$ constraints. The covariance matrix of these constrained parameter estimates is derived in the following way. Let $\hat{\mathbf{k}}$ denote the $2D$ vector with constraints. The first order partial derivatives of $\hat{\mathbf{k}}$ with respect to the P^c parameters are collected in the $2D$ by P^c matrix $\hat{\mathbf{K}}$. Let $\hat{\mathbf{A}} = \hat{\mathbf{G}}'^c\hat{\mathbf{G}}^c + \hat{\mathbf{K}}'\hat{\mathbf{K}}$, where $\hat{\mathbf{G}}^c$ is the NT by P^c matrix with first order partial derivatives of the model with respect to the P^c parameters. The covariance matrix of the constrained parameter estimates then is (Browne & du Toit, 1992; Huizenga & Molenaar, 1994; Schott, 1997, p. 248)

$$\hat{\mathbf{C}}^c = \hat{\sigma}^2[\hat{\mathbf{A}}^{-1} - \hat{\mathbf{A}}^{-1}\hat{\mathbf{K}}'[\hat{\mathbf{K}}\hat{\mathbf{A}}^{-1}\hat{\mathbf{K}}']^{-1}\hat{\mathbf{K}}\hat{\mathbf{A}}^{-1}]. \quad (5)$$

$\hat{\sigma}^2$ is now estimated by $s(\hat{\boldsymbol{\theta}}^c)/(NT - P^c + 2D)$. $\hat{\mathbf{C}}^c$ has $P^c - 2D$ positive eigenvalues, and $2D$ eigenvalues equal to zero. Tests on these constrained source parameters can easily be derived by generalizing (4) to the constrained case.

Increasing Accuracy

The covariance matrix (3) depends on the noise variance σ^2 , and thus on the number of trials M . Moreover, it depends on the derivatives \mathbf{G} , and thus on the number of samples T , and on the number and position of sensors \mathbf{X} . Our objective is to determine the number of trials, the number of samples, and the number and position of sensors required for a covariance matrix with the desired properties. Since the covariance matrix depends on the actual values of the source parameters, the method should be based on an explicit source hypothesis. This hypothesis can be derived from anatomical knowledge, from animal studies, or from Positron Emission Tomography or functional Magnetic Resonance Imaging studies. It can also be derived from previously reported EEG/MEG source studies, or from a pilot study, as will be illustrated later.

The sensor configuration generating an optimal covariance matrix is determined by minimizing the generalized variance of the parameters, as indexed by the product of the eigenvalues of \mathbf{C} or the determinant $|\mathbf{C}|$ (Atkinson & Donev, 1992). Equivalently, $|\mathbf{G}'\mathbf{G}|$ can be maximized since σ^2 is independent of sensors and samples under the assumption of homoscedastic and uncorrelated noise. This so called D -optimality criterion was used for example by Hochwald and Nehorai (1997) to determine the optimal orientation of MEG sensors.

After an optimal sensor configuration is calculated, it should be determined whether this sensor configuration yields the required accuracy. This can be assessed by computing confidence regions given the source hypothesis and given an initial guess of the noise variance (Huizenga & Molenaar, 1994). Therefore these confidence regions will be called a priori confidence regions. Alternatively, if an optimal configuration is computed for several sources, then (4) can be used

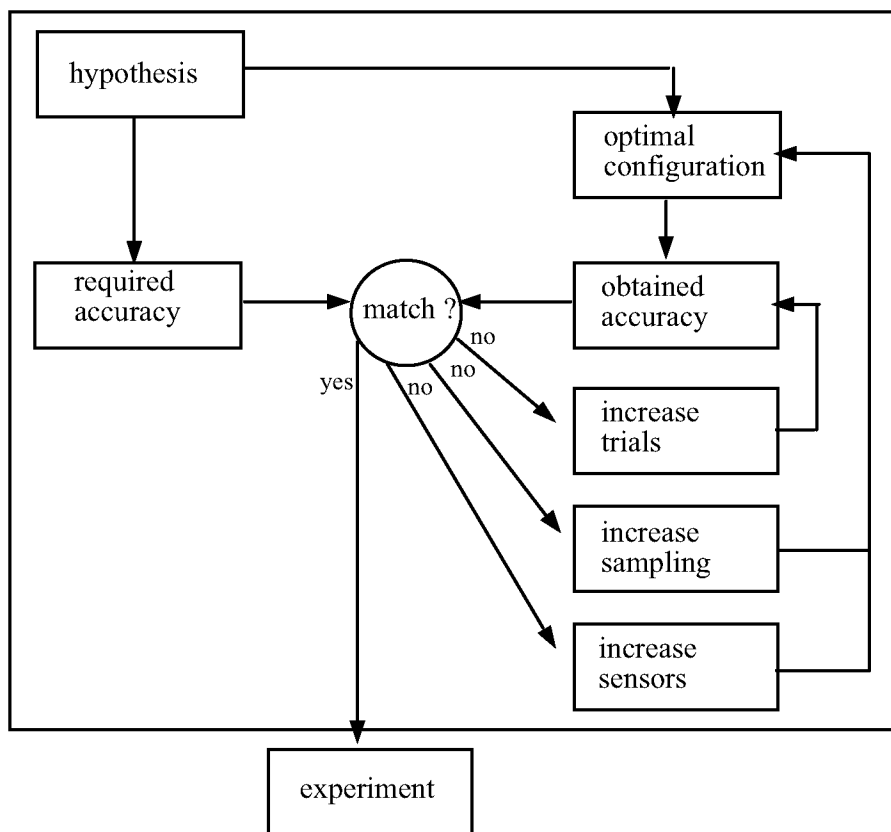


FIGURE 1.
Flowchart to determine the optimal measurement conditions.

to test whether these sources differ significantly given the source hypothesis and given an initial guess of the noise variance. Therefore these tests will be called *a priori* tests.

If the optimal sensor configuration does not yield required accuracy, then there are several options to increase accuracy (Figure 1). First, the noise variance can be decreased by increasing the number of trials. Let σ_s^2 be the variance of a trial, then $\sigma_k^2 = \sigma_s^2/k$ and $\sigma_l^2 = \sigma_s^2/l$ denote the variance of the mean of k and l trials respectively. If a known σ_k^2 is too high, and should be replaced by a lower noise variance σ_l^2 , then the required number of trials l is derived from $l = k\sigma_k^2/\sigma_l^2$. Second, precision can be increased by increasing the number of sensors. This necessitates repeated cycles of adding a set of sensors and renewed optimization of the sensor configuration (Figure 1). Third, precision can also be improved by increasing the number of samples. In this case it is also required to perform repeated cycles of adding a set of samples, and renewed optimization of the sensor configuration (Figure 1). Since, as will be shown later, the optimal sensor configuration is dependent on the number of samples.

Optimization of the D -Optimality Criterion

We describe two routines to attain D -optimality. The continuous optimizer allows for sensors around the entire scalp. The discrete optimizer chooses among predetermined positions, for example in an electrocap. In each case we first describe optimization of the unconstrained covariance matrix \mathbf{C} , and then proceed with the constrained matrix \mathbf{C}^c .

If sensors are allowed on the entire scalp, then the problem is continuous. In this case the generalized variance $|\mathbf{C}|$ can be minimized by standard nonlinear optimization routines. We use the quasi-Newton algorithm with finite difference gradients. If the constrained parametrization is used, then the generalized variance in $P^c - 2D$ subspace should be minimized, as indexed by the product of the $P^c - 2D$ nonzero eigenvalues of \mathbf{C}^c (see Pukelsheim, 1993).

If sensors are restricted to predetermined positions, then the design problem is discrete. It consists of choosing a fixed number of N sensors out of L candidate sensors to maximize $|\mathbf{G}'\mathbf{G}|$. A naive approach is to compute $|\mathbf{G}'\mathbf{G}|$ for $L!/N!(L - N)!$ configurations, and then select the optimal configuration. However, even a limited task of selecting $N = 30$ out of $L = 60$ sensors, leads to huge computational difficulties: calculation and comparison of approximately 10^{16} configurations. Therefore a more efficient technique is required, for example the exchange algorithm developed by Fedorov (see Atkinson & Donev, 1992; St. John & Draper, 1975). The standard exchange algorithm is only suited for a single response, that is instantaneous, analysis. Therefore we describe a generalization which is also applicable to the multiresponse spatiotemporal case.

In the exchange algorithm a sensor from a starting configuration is exchanged for a candidate sensor if this exchanging improves $|\mathbf{G}'\mathbf{G}|$ to a maximal extent. This process is repeated until no further increase in the determinant can be attained. If a sensor i is added to a configuration, and a sensor j is removed, then $\mathbf{G}'\mathbf{G}$ becomes $[\mathbf{G}'\mathbf{G} - \mathbf{U}_{ij}\mathbf{V}'_{ij}]$, with P by $2T$ matrices:

$$\begin{aligned}\mathbf{U}_{ij} &= [-\mathbf{g}_{i1}, -\mathbf{g}_{i2}, \dots, -\mathbf{g}_{iT}, \mathbf{g}_{j1}, \mathbf{g}_{j2}, \dots, \mathbf{g}_{jT}] \\ \mathbf{V}_{ij} &= [\mathbf{g}_{i1}, \mathbf{g}_{i2}, \dots, \mathbf{g}_{iT}, \mathbf{g}_{j1}, \mathbf{g}_{j2}, \dots, \mathbf{g}_{jT}],\end{aligned}\quad (6)$$

where \mathbf{g}_{it} is the P vector with first order partial derivatives of the model at sensor i and sample t . Sensors i and j should be chosen such that the increase in determinant is maximal. If $\mathbf{G}'\mathbf{G}$ is nonsingular then $|\mathbf{I}[\mathbf{G}'\mathbf{G} - \mathbf{U}_{ij}\mathbf{V}'_{ij}]| = |\mathbf{G}'\mathbf{G}| |\mathbf{I} - \mathbf{V}'_{ij}[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{U}_{ij}|$, where \mathbf{I} is the $2T$ by $2T$ unity matrix (Schott, 1997, p. 250). Therefore, the following criterion can be maximized:

$$\frac{|\mathbf{G}'\mathbf{G} - \mathbf{U}_{ij}\mathbf{V}'_{ij}|}{|\mathbf{G}'\mathbf{G}|} = \frac{|\mathbf{G}'\mathbf{G}| |\mathbf{I} - \mathbf{V}'_{ij}[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{U}_{ij}|}{|\mathbf{G}'\mathbf{G}|} = |\mathbf{I} - \mathbf{V}'_{ij}[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{U}_{ij}|. \quad (7)$$

In the first iteration Equation (7) is evaluated for $N(L - N)$ exchanges and the best exchange is chosen. Then $[\mathbf{G}'\mathbf{G}]^{-1}$ is updated with the chosen exchange. This exchanging and updating pro-

cedure is iterated until no further increment in the determinant can be found. In the instantaneous case, $T = 1$, and (7) reduces to

$$1 + \mathbf{g}'_i [\mathbf{G}'\mathbf{G}]^{-1} \mathbf{g}_i - \mathbf{g}'_j [\mathbf{G}'\mathbf{G}]^{-1} \mathbf{g}_j - \mathbf{g}'_i [\mathbf{G}'\mathbf{G}]^{-1} \mathbf{g}_j \mathbf{g}'_j [\mathbf{G}'\mathbf{G}]^{-1} \mathbf{g}_j + \mathbf{g}'_j [\mathbf{G}'\mathbf{G}]^{-1} \mathbf{g}_i \mathbf{g}'_i [\mathbf{G}'\mathbf{G}]^{-1} \mathbf{g}_j. \quad (8)$$

Which is the standard single response Fedorov exchange criterion (e.g., Cook & Nachtsheim, 1980).

In the constrained case, the discrete optimizer proceeds in an equivalent way. However, since $\mathbf{G}'^c \mathbf{G}^c$ is singular, it is replaced by the nonsingular matrix \mathbf{A} from (5); refer to Browne & Du Toit, 1992; or Silvey, 1970, p. 63. In the simulation section it is shown that this yields acceptable results.

Instantaneous Versus Spatiotemporal Optimality

Suppose $|\mathbf{G}'\mathbf{G}|$ can be written as a sum or product of two functions, where one function is independent of sensor positions, and the other function is independent of T and Ψ . In this case an instantaneous optimal configuration is also optimal for spatiotemporal data and vice versa. It then suffices to optimize the single response instantaneous criterion. However, this is not the case, as is shown below.

The partial derivative matrix \mathbf{G} is partitioned in a TN by $5D$ matrix related to the $5D$ fixed location and orientation parameters, and in a TN by TD matrix related to the TD varying amplitude parameters. \mathbf{G} equals

$$\mathbf{G} = [(\Psi' \otimes \mathbf{I}_N)\mathbf{Z}; (\mathbf{I}_T \otimes \mathbf{H})], \quad (9)$$

where \mathbf{I}_N and \mathbf{I}_T are unit matrices of size N by N and T by T respectively. \mathbf{Z} is the ND by $5D$ matrix

$$\mathbf{Z} = \begin{bmatrix} \partial \mathbf{H}_1 & 0 & 0 & 0 \\ 0 & \partial \mathbf{H}_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \partial \mathbf{H}_D \end{bmatrix}. \quad (10)$$

Where each N by 5 matrix $\partial \mathbf{H}_d$ contains the partial derivatives of the lead field of source d with respect to its 5 location and orientation parameters. $\mathbf{G}'\mathbf{G}$ then equals

$$\mathbf{G}'\mathbf{G} = \begin{bmatrix} \mathbf{Z}'(\Psi \otimes \mathbf{I}_N)(\Psi' \otimes \mathbf{I}_N)\mathbf{Z} & \mathbf{Z}'(\Psi \otimes \mathbf{I}_N)(\mathbf{I}_T \otimes \mathbf{H}) \\ (\mathbf{I}_T \otimes \mathbf{H}')(\Psi' \otimes \mathbf{I}_N)\mathbf{Z} & (\mathbf{I}_T \otimes \mathbf{H}'\mathbf{H}) \end{bmatrix}. \quad (11)$$

Let $[\mathbf{I}_N - \mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'] = \mathbf{B}$, then (Schott, 1997, p. 250, 256)

$$|\mathbf{G}'\mathbf{G}| = |\mathbf{H}'\mathbf{H}|^T |\mathbf{Z}'(\Psi\Psi' \otimes \mathbf{B})\mathbf{Z}|. \quad (12)$$

The source amplitudes can be isolated from the second term if $\Psi\Psi'$ is diagonal, then:

$$|\mathbf{G}'\mathbf{G}| = |\mathbf{H}'\mathbf{H}|^T \left(\prod_{d=1}^D (\Psi'_d \Psi_d)^5 \right) \left(\prod_{d=1}^D |\partial \mathbf{H}'_d \mathbf{B} \partial \mathbf{H}_d| \right) \quad (13)$$

where Ψ_d contains the T amplitudes of source d . The second term in (13) is independent of sensors, and the third term is independent of the number of samples and independent of the source amplitudes. However, the first term still is a function of both the sensors and the number of samples. Therefore, instantaneous and spatiotemporal optimal configurations are not equivalent.

Simulations

The purpose of the simulations is to (a) get an indication of the improvements obtained by using continuously or discretely optimized configurations, (b) investigate the constrained criterion, (c) investigate the use of a source hypothesis derived from a pilot, and (d) investigate the benefits of an optimal configuration in case of multiple sources. These issues can be addressed in simulations with the instantaneous model. Differences between instantaneous and spatiotemporal D-optimality are investigated in the final simulation section.

Optimal sensor configurations are derived by either the continuous or discrete optimizer. The optimizers are restarted 5 times with different starting values to prevent local minima. The data are simulated on a standard configuration with 41 sensors uniformly distributed around the scalp; and on an optimal set of 41 sensors. The data are simulated with respect to a hypothetical zero reference. Since the simulations are very time consuming, they only concern one or two sources in the computationally simple homogenous sphere with a radius $r = 10.00$ cm. Normally distributed homoscedastic and uncorrelated noise with zero mean is added to the signals, therefore sources are estimated by OLS. The source parameters in the normalized Euclidean parametrization are estimated by the augmented Lagrangian algorithm, which takes the constraints on these parameters into account (e.g., Gill, Murray, & Wright, 1981).

In order to get an indication of precision, the simulations are performed in 400 replicates. Several measures are reported. The a priori estimated standard errors are the standard errors given an a priori hypothesis on the source parameters and the noise variance. The simulation standard deviations are the standard deviations of the parameter estimates across the 400 replicates. The reported standard errors and standard deviations are multiplied by r , the radius of the sphere. The a priori improvement denotes the ratio of the a priori estimated standard errors on optimized and standard configurations, averaged over parameters. The simulation improvement denotes the ratio of the simulation standard deviations on optimized and standard configurations, averaged over parameters. Finally, the average Euclidean distance between true and estimated vectors is reported, it equals

$$\text{location error: } \frac{r}{400} \sum_{j=1}^{400} \left((\hat{\varphi}_j \hat{\tau}_{1j}^n - \varphi \tau_1^n)^2 + (\hat{\varphi}_j \hat{\tau}_{2j}^n - \varphi \tau_2^n)^2 + (\hat{\varphi}_j \hat{\tau}_{3j}^n - \varphi \tau_3^n)^2 \right)^{1/2}$$

$$\text{moment error: } \frac{r}{400} \sum_{j=1}^{400} \left((\hat{\psi}_j \hat{\xi}_{1j}^n - \psi \xi_1^n)^2 + (\hat{\psi}_j \hat{\xi}_{2j}^n - \psi \xi_2^n)^2 + (\hat{\psi}_j \hat{\xi}_{3j}^n - \psi \xi_3^n)^2 \right)^{1/2},$$

where $\hat{\tau}_{1j}^n, \hat{\tau}_{2j}^n, \dots$ denote the estimates in 400 simulation replicates, and $\tau_1^n, \tau_2^n, \dots$ denote the true parameters.

One Source, Instantaneous Criterion

We calculated optimal configurations for a single source defined by $\tau_1^n = -.56$, $\tau_2^n = .20$, $\tau_3^n = .80$, $\xi_1^n = .56$, $\xi_2^n = -.20$, $\xi_3^n = .80$, $\varphi = .75$, $\psi = .40$. The noise variance was set to a realistic value, yielding a signal to noise ratio of 7.10dB on the standard configuration.

First, we optimized precision given a true source hypothesis, and by continuously minimizing the product of the nonzero eigenvalues of \mathbf{C}^c in (5). In Figure 2 it can be seen that the optimal configuration favors the region with the highest spatial gradient around the polarity reversal. The optimal configuration yielded a considerable improvement: the location and moment errors, the a priori estimated standard errors, and the simulation standard deviations decreased to a large extent (Figure 3, Table 1, compare first and second row). Figure 3 shows in addition that the a priori estimated standard errors predicted the simulation standard deviations very well. The latter result is important since it is an essential prerequisite for the proposed methodology.

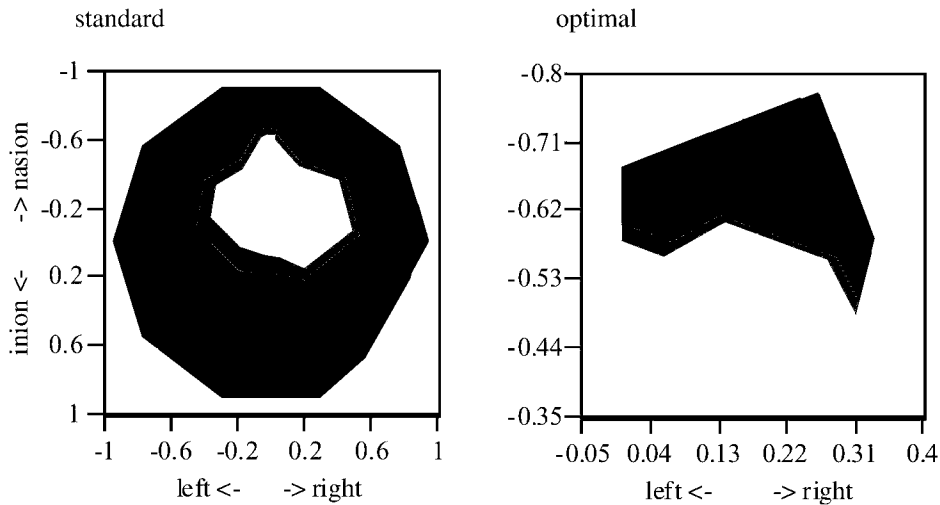


FIGURE 2. The isocontours of the potential field (in arbitrary units) on standard and optimal configurations. The latter was derived by continuously optimizing \mathbf{C}^c given the true one source hypothesis. The coordinates are given in the unit sphere.

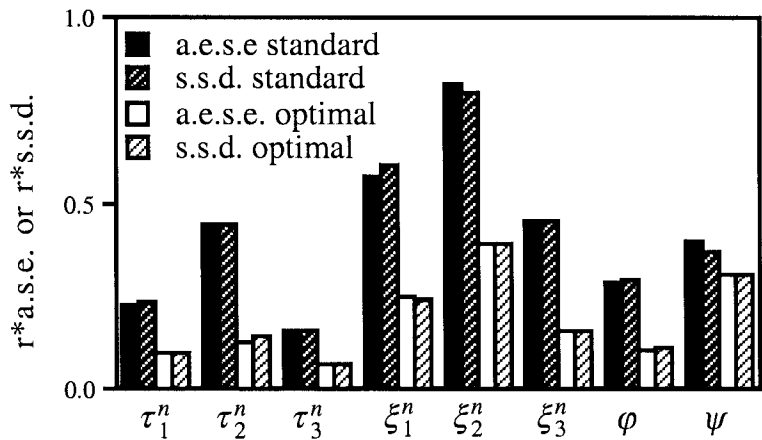


FIGURE 3. A priori estimated standard errors multiplied by the radius of the sphere ($r^*a.s.e.$) and simulation standard deviations multiplied by the radius of the sphere ($r^*s.s.d.$) for standard and optimal configurations. The latter was derived by continuously optimizing \mathbf{C}^c given the true one source hypothesis.

TABLE 1. Standard and optimal configurations for one source

Configuration	Location error	Moment error	A priori improvement	Simulation improvement
standard	0.44	0.52	1.00	1.00
continuous, true, \mathbf{C}^c	0.16	0.33	0.41	0.42
continuous, true, \mathbf{A}	0.17	0.34	0.45	0.45
discrete, true, \mathbf{A}	0.33	0.41	0.74	0.74
continuous, pilot, \mathbf{C}^c	0.16	0.30	0.42	0.43

Note: The optimal configurations are differentiated by: (a) continuous or discrete optimization; (b) true or pilot hypothesis; and (c) optimization of the product of nonzero eigenvalues of \mathbf{C}^c or the determinant of \mathbf{A} .

Second, we investigated given a true source hypothesis, whether continuously maximizing $|\mathbf{A}|$ is comparable to continuously minimizing the product of the nonzero eigenvalues of \mathbf{C}^c (see (5)). The results in Table 1 (compare second and third row) indicate that this is the case. This is a favorable result that supports optimization of $|\mathbf{A}|$ in case of a discrete configuration.

Third, we derived a discrete optimal configuration given a true source hypothesis. The discrete optimizer maximized $|\mathbf{A}|$ by choosing 41 sensors out of a candidate set of 105 sensors, which consisted of the aforementioned 41 standard sensors plus 64 additional sensors. In Table 1 (fourth row) it can be seen that the discrete optimizer improves precision, although self evidently less than the continuous optimizer.

Fourth, we investigated whether the method is also valid if it is based on a source hypothesis derived from a pilot. We selected, from the distribution of 400 estimates on the standard configuration, an estimate that was located at a maximal distance (0.72 cm) from the true source. Subsequently, we continuously optimized the sensor configuration for this estimated source. We then performed simulations for this optimal configuration and the original true source. In Table 1 (fifth row) it is shown that all results are comparable to the configuration computed from a true source hypothesis. This indicates that although a pilot may be sub optimal, it may guide the choice of a sensor configuration for an optimal experiment.

In sum, an optimal sensor configuration offers a notable accuracy improvement. This optimal sensor configuration can either be computed by continuously minimizing the nonzero eigenvalues of \mathbf{C}^c or by continuously maximizing $|\mathbf{A}|$. Discrete maximization of $|\mathbf{A}|$ also increases accuracy. Finally, these optimal configurations can be based on a source hypothesis derived from a pilot.

Two Sources, Instantaneous Criterion

If a sensor configuration does not provide enough information to separate sources, then the source estimates are very unreliable (Huizenga, 1995, chap. 3). Therefore it is essential to determine a priori whether the configuration contains enough information to separate the sources, and to optimize the configuration if this is not the case.

We applied this method to two sources separated 3.00 cm, and mirroring each other in the midline. Only τ_2^n and ξ_2^n differ between the two sources, the parameters are: $\tau_1^n = .20$; $\tau_2^n = .20, -.20$; $\tau_3^n = .96$; $\xi_1^n = .20$; $\xi_2^n = .20, -.20$; $\xi_3^n = .96$; $\varphi = .75$ and $\psi = .40$. Noise variance was chosen to get a realistic signal to noise ratio of 7.10dB on the standard configuration. The data were modeled by two or one source(s). If the sensor configuration does not provide enough information to separate the sources, then the a priori test on τ_2^n will be nonsignificant. Moreover, the Lack of Fit test (Huizenga & Molenaar, 1994) then will indicate that there is no lack of fit for the single source model.

The a priori test on τ_2^n indicated that the source locations did not differ significantly on the standard configuration ($F = .95$, $df = 1, 29$, $p = .95$). This a priori test was indicative of the simulation results: location and moment errors of the two sources were very large (Table 2, first

TABLE 2.
Standard and optimal configuration for two sources

Configuration	D	% LoF	Location error	Moment error
Standard	2	3%	3.76/3.41	5.52/5.38
	1	14%	1.70	4.25
Optimal	2	3%	.86/.93	3.99/4.00
	1	95%	2.11	8.97

Note: “ D ” is the number of estimated sources, and “% LoF” is the percentage of simulation replicates with a Lack of Fit. If one source is estimated, then the location and moment errors refer to the error between the estimated and the nearest true source.

row), and a single source did fit the data, only in 14% of the replicates there was a lack of fit (Table 2, second row).

Results improved on the optimal configuration. First, the a priori test was significant ($F = 28.62$, $df = 1, 29$, $p < .001$). Second, the location and moment errors of the two sources decreased (Table 2, third row). Third, a single source did not fit the data: in 95% of the replicates there was a lack of fit (Table 2, fourth row). For these reasons, the optimal sensor configuration seems to yield enough information to separate the two sources.

Two Sources: Instantaneous vs. Spatiotemporal D-Optimality

An instantaneous optimal design is suboptimal for spatiotemporal data. In order to investigate the degree of suboptimality, the discrete optimizer chose 41 out of 105 sensors, according to the following two criteria: (i) optimal for an entire time window, or (ii) optimal for the first sample only. Subsequently, the a priori estimated standard errors for the parameters in the spatiotemporal model were computed given the standard, instantaneous optimal, and spatiotemporal optimal configurations.

The optimal configurations were computed for two simultaneously active sources with parameters $\tau_1^n = .10$; $\tau_2^n = .70, -.70$; $\tau_3^n = .71$; $\xi_1^n = .10$; $\xi_2^n = .10$; $\xi_3^n = .99$; $\varphi = .75$. Three cases were investigated. In the first case, $\Psi\Psi'$ was diagonal and the time window consisted of 10 samples (source I amplitude was always .4; source II amplitude was .4 in the first 5 samples and $-.4$ in the second 5 samples). In the second case, $\Psi\Psi'$ was also diagonal, but there were 20 samples (source I amplitude was always .4; source II amplitude was .4 in the first 10 samples, and $-.4$ in the second 10 samples). In the third case, $\Psi\Psi'$ was not diagonal and the time window consisted of 10 samples (source I amplitude was always .4; source II amplitude was .4, .4, .4, .4, .4, .4, .4, .4, .0, $-.46$).

In all cases the instantaneous optimal configuration was suboptimal for spatiotemporal data: the amplitude parameters were less precise, whereas the location and orientation parameters were more precise (Figure 4). Note however that the differences were very small. Similar small differences between spatiotemporal and instantaneous optimal configurations were found for a variety of sources and a variety of amplitude functions.

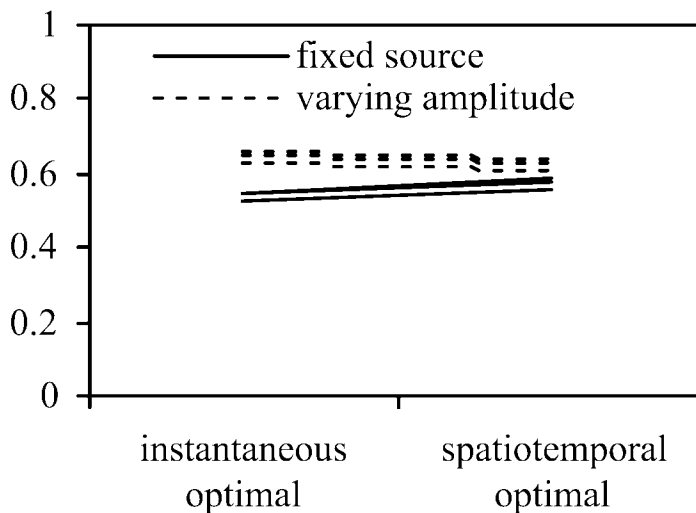


FIGURE 4.

A priori improvement for the spatiotemporal model given an instantaneous or spatiotemporal optimal configuration. The improvement is given separately for fixed location and orientation parameters; and time varying amplitude parameters.

Empirical Illustration

In this section we give an empirical illustration of an optimal configuration derived from a pilot experiment. A pilot can be carried out on one subject with a standard sensor configuration. The estimated sources of this pilot then may generate the source hypothesis for the optimal configuration to be used with the entire subject sample. However, this setup is not convenient for our present methodological purposes. For a lowering of estimated standard errors can also be introduced by a lack of replicability between pilot and optimal experiments. More specifically, the noise variance may decrease, or the sources may differ. Therefore we conceived a setup in which the pilot and optimal experiment are measured at once, ensuring replicability between the two runs.

Data were recorded from 64 sensors covering the entire head. First we analyzed a subset, namely 32 sensors approximately equally spaced and covering the entire head. Based on the source estimate of this standard subset, we calculated an optimal 32 subset. This subset was computed under the assumption that the noise was homoscedastic and uncorrelated, since we had no a priori hypothesis concerning the noise structure. We then estimated the sources on the optimal subset. Finally, we estimated the sources of the entire set of 64 sensors. The latter analysis served as a gold standard generating the most adequate estimates. If the proposed methodology is valid, then the standard errors should decrease on the optimal compared to the standard subset. In addition the parameter and standard error estimates derived from 64 sensors should be approximated better by the optimal subset than by the standard subset.

Method

The subject was one right-handed male, 23 years of age, with normal vision. Stimuli were presented equiprobably and randomly to the lower left or lower right quadrant of the visual field, at a radial eccentricity of 4° , with an inter-stimulus interval varying between 550 and 750 ms. Stimuli were vertical, white bars, flashed for 50 ms on a dark background. The height of the bars was either 2.4° in 86% of the trials, or 3.2° in 14% of the trials. The task was to respond manually to the longer, infrequent bars in either the right or left quadrant, while fixating the center of the screen. Here we focus on 4 samples of the P1 component (120.0 – 127.5 ms) evoked by attended non-target stimuli (shorter bars) in the lower right quadrant. The data were filtered and trials contaminated by eye artifacts were removed, leaving 213 trials for analysis. EEG was recorded from 64 Ag/AgCl sensors mounted in an elastic cap, referenced to the left mastoid. A sphere was fitted through the sensors, the estimated sphere radius was $r = 9.48$ cm. The sensors were radially projected on this sphere, the resulting sensor positions are depicted in Figure 5.

In an instantaneous analysis, the source was modeled as one dipole in three isotropic concentric spheres with radii of 87%, 92% and 100% of $r = 9.48$ cm, and conductivities of 1.0, .0125 and 1.0. In the second panel of Figure 5 it can be seen that noise, that is the trial variation around the mean, was heteroscedastic and correlated. This was confirmed by appropriate statistical tests (Huizenga & Molenaar, 1995). Therefore the source parameters and their standard errors were estimated by generalized least squares (GLS) (Huizenga & Molenaar, 1996).

Empirical Results

In Figure 5 we have depicted sensor configurations, noise characteristics, parameter estimates and estimated standard errors. Since the samples were homogeneous with respect to all of these indices, the results refer to averages over samples. The average noise variance was equal on the standard and optimal subset. This excludes the possibility that the estimated standard errors decrease due to a lowering of the noise variance. All estimated standard errors were low on the optimal compared to the standard subset, although the standard error of ξ_1^n was higher. This may be due to the fact that the precision of all parameters is optimized, which may be at the cost of parameters that are already precise, like ξ_1^n . Finally, the parameter estimates of the optimal sub-

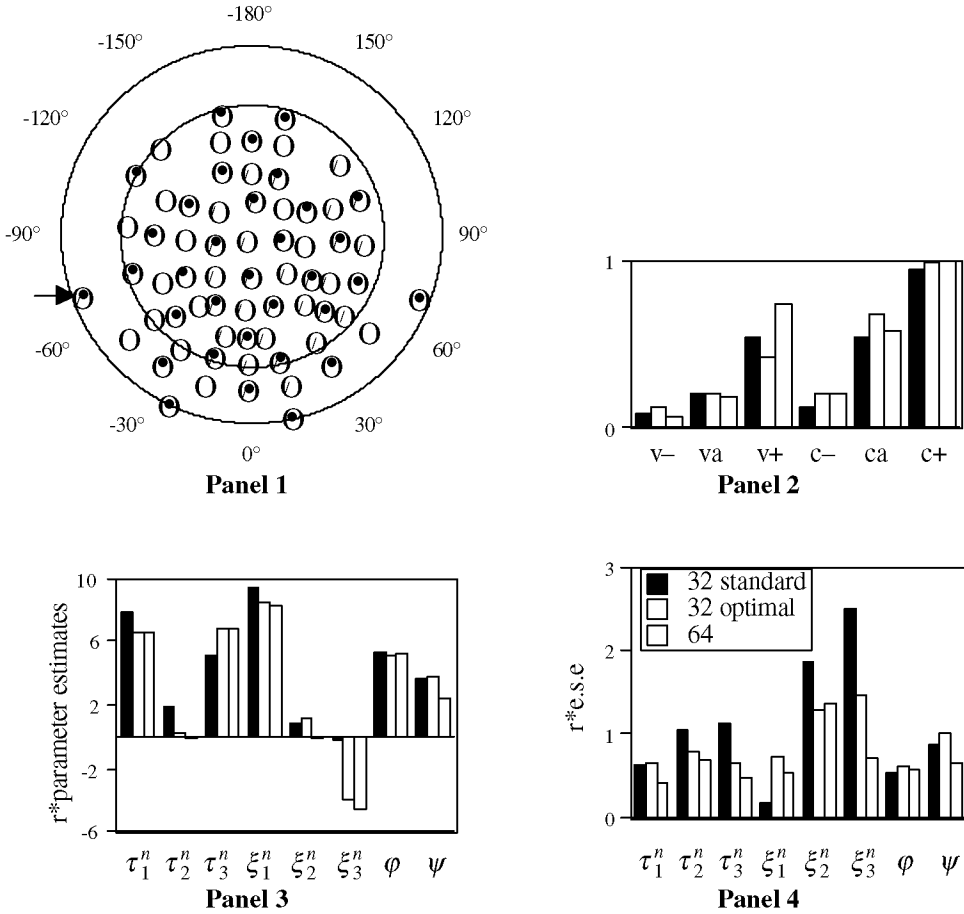


FIGURE 5.

In the upper panel the empirical sensor configurations. “O”: the entire set of 64 sensors; “•”: the standard 32-subset; “/” the optimal 32-subset. The arrow points towards the reference, which was equal for all configurations. The plot is given in the spherical coordinates latitude ($-180^\circ - 180^\circ$) and longitude ($0^\circ - 180^\circ$), the two circles refer to a longitude of 90° and 130° . The second panel contains noise characteristics: v-, va, v+ denote minimal, average and maximal noise variance; c-, ca, c+ denote minimal, average, and maximal noise correlation. The third panel contains the parameter estimates multiplied by the radius of the sphere ($r \cdot \text{parameter estimates}$). The fourth panel contains the estimated standard errors multiplied by the radius of the sphere ($r \cdot \text{e.s.e.}$).

set approximated the parameter estimates of the entire set of 64 sensors. Moreover, the standard errors of the optimal subset approximated the standard errors of the entire set.

In sum, the optimal subset yields more precise estimates than the standard subset: the standard errors are smaller, and the parameter and standard error estimates of the optimal subset approximated those of the entire set.

Discussion

The method summarized in Figure 1 has several advantages. First, it is analytic and therefore fast. For example, if the single response discrete optimizer is used, then it takes only a few seconds to choose 32 sensors out of a set of 64 sensors. Computational demands will increase for more realistic head models, but the demands can be kept low by computing model derivatives in advance for all possible sensors. A second advantage is that the method improves accuracy considerably. The improvement depends on the source configuration under study, so it is difficult

to indicate to what extent the estimates will generally ameliorate. But to give an indication, in the simulations the standard errors halved if sensors were positioned optimally. Moreover, it became possible to separate two simultaneously active sources, which was impossible on a standard configuration. A third advantage is that experimental demands are optimized at minimal costs. Positioning sensors optimally does not cost anything. Moreover, the required number of sensors, samples and trials can be calculated explicitly, thereby preventing experiments with unnecessarily high demands. A fourth advantage is that the method may indicate a priori that experimental requirements are not attainable. This offers the possibility to prevent experiments with insufficient power.

It was shown that instantaneous and spatiotemporal optimal configurations are not equivalent. Therefore, the multiresponse Fedorov exchange should be used for spatiotemporal data. The differences are small however, and therefore the standard single response exchange may be used if computational time is really an issue.

The proposed methods are based on the a priori covariance matrix of the parameter estimates. Therefore this matrix should be accurate, necessitating three requirements. First, it is necessary that the regression model is approximately linear around the parameter estimates. If this is not the case, then optimal configurations based on likelihood approaches may be more appropriate (Seber & Wild, 1989). A second prerequisite is that the noise is homoscedastic and uncorrelated, both in space and in time. This requirement can be avoided by incorporating a valid a priori hypothesis on the noise covariance matrix (Huizenga, De Munck, Waldorp, & Grasman, in press; Waldorp, Huizenga, Dolan, & Molenaar, 2001). A third requirement is a good source hypothesis. The method seems to be quite robust against moderate errors in the source hypothesis, since a hypothesis derived from a pilot yielded a nearly optimal sensor configuration. However, the method will fail if the hypothesis can not be formulated reliably. In that case it is advised to measure the maximum number of samples, trials and sensors, and to distribute the sensors uniformly.

The results of the empirical study indicate that a pilot may guide the choice of experimental requirements. Two reservations should be made however. First, the pilot and the optimal configuration were measured simultaneously on one subject. In actual applications, a pilot will be conducted first. Therefore, low replicability may lower the gain of the proposed methodology. Second, we based our optimal configuration on the assumption that noise was homoscedastic and uncorrelated. This assumption is however in general not satisfied. Therefore, even better results can be found by incorporating a reliable a priori noise model (Huizenga et al., in press; Waldorp et al., 2001).

It is worthwhile to mention two extensions. First, sometimes a small rigid MEG sensor array is used (e.g., Gunji, Kagigi & Hoshiyama, 2000). The optimal placement of this array can be determined by extending the present method to incorporate rigid constraints on the sensor positions. The second extension concerns simultaneous EEG and MEG analysis. It has been shown that adding a few EEG sensors to a MEG configuration is very beneficial (Huizenga, van Zuijen, Heslenfeld, & Molenaar, 2001). The present method can be extended to determine the optimal positioning of these EEG sensors, given a fixed MEG configuration.

References

- Ahonen, A.I., Hämläinen, M.S., Ilmoniemi, R.J., Kajola, M.J., Knuutila, J.E.T., Simola, J.T., & Vilkman, V.A. (1993). Sampling theory for neuromagnetic detector arrays. *IEEE Transactions on Biomedical Engineering*, 40(9), 859–869.
- Atkinson, A.C., & Donev, A.N. (1992). *Optimum experimental designs*. Oxford, U.K.: Clarendon Press.
- Browne, M.W., & du Toit, S.H.C. (1992). Automated fitting of nonstandard models. *Multivariate Behavioral Research*, 27, 269–300.
- Cook, R.D., & Nachtsheim, C.J. (1980). A comparison of algorithms for constructing exact D-optimal designs. *Technometrics*, 22(3), 315–324.
- Cuffin, B.N. (1985). A comparison of moving dipole inverse solutions using EEG's and MEG's. *IEEE Transactions on Biomedical Engineering*, 32(11), 905–910.
- Cuffin, B.N., & Cohen, D. (1979). Comparison of the magnetoencephalogram and electroencephalogram. *Electroencephalography and Clinical Neurophysiology*, 47, 132–146.

- Gaumond, R.P., Lin, J.-H., & Geselowitz, D.B. (1983). Accuracy of dipole localization with a spherical homogeneous model. *IEEE Transactions on Biomedical Engineering*, 30(1), 29–34.
- Gill, P.E., Murray, W., & Wright, M.H. (1981). *Practical Optimization*. London, U.K.: Academic Press.
- Gunji, A., Kagigi, R., & Hoshiyama, M. (2000). Spatiotemporal source analysis of vocalization-associated magnetic fields. *Cognitive Brain Research*, 9, 157–163.
- Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., & Lounasmaa, O.V. (1993). Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65, 413–497.
- Hari, R., Joutsiniemi, S.-L., & Sarvas, J. (1988). Spatial resolution of neuromagnetic records: Theoretical calculations in a spherical model. *Electroencephalography and Clinical Neurophysiology*, 71, 64–72.
- Hochwald, B., & Nehorai, A. (1997). Magnetoencephalography with diversely oriented and multicomponent sensors. *IEEE Transactions on Biomedical Engineering*, 44, 40–50.
- Huizenga, H.M. (1995). *The statistical approach to electromagnetic source localization in the brain*. Unpublished doctoral dissertation, University of Amsterdam, Amsterdam.
- Huizenga, H.M., De Munck, J.C., Waldorp, L.J., & Grasman, R.P.P. (in press). Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model. *IEEE Transactions on Biomedical Engineering*.
- Huizenga, H.M., & Molenaar, P.C.M. (1994). Estimating and testing the sources of evoked potentials in the brain. *Multivariate Behavioral Research*, 29, 237–262.
- Huizenga, H.M., & Molenaar, P.C.M. (1995). Equivalent source estimation of scalp potential fields contaminated by heteroscedastic and correlated noise. *Brain Topography*, 8, 13–33.
- Huizenga, H.M., & Molenaar, P.C.M. (1996). Ordinary least squares dipole localization is influenced by the reference. *Electroencephalography and Clinical Neurophysiology*, 99, 562–567.
- Huizenga, H.M., van Zuijlen, T.L., Heslenfeld, D.J., & Molenaar, P.C.M. (2001). Simultaneous MEG and EEG source analysis. *Physics in Medicine and Biology*, 46(7), 1737–1751.
- Kenemans, J.L., Baas, J.M.P., Mangun, G.R., Lijffijt, M., & Verbaten, M.N. (2000). On the processing of spatial frequencies as revealed by evoked-potential source modeling. *Clinical Neurophysiology*, 111, 1113–1123.
- Kuc, R. (1996). Magnetometer spacing criterion for biomagnetic source current imaging. *IEEE Transactions on Biomedical Engineering*, 43(11), 1125–1127.
- Lange, N., & Zeger, S.L. (1997). Non-linear Fourier time series analysis for human brain mapping by functional magnetic resonance imaging. *Applied Statistics*, 46, 1–29.
- Mosher, J.C., Spencer, M.E., Leahy, R.M., & Lewis, P.S. (1993). Error bounds for EEG and MEG dipole source localization. *Electroencephalography and Clinical Neurophysiology*, 86, 303–321.
- Nunez, P.L. (1988). Spatial filtering and experimental strategies in EEG. In D. Samson-Dollfus (Ed.), *Statistics and topography in quantitative EEG*. Paris, France: Elsevier.
- Ogura, Y., & Sekihara, K. (1993). Relationship between dipole parameter estimation errors and measurement conditions in magnetoencephalography. *IEEE Transactions on Biomedical Engineering*, 40(9), 919–924.
- Pukelsheim, F. (1993). *Optimal design of experiments*. New York, NY: Wiley.
- Schott, J.R. (1997). *Matrix analysis for statistics*. New York, NY: Wiley.
- Seber, G.A.F., & Wild, C.J. (1989). *Nonlinear regression*. New York, NY: Wiley.
- Silvey, S.D. (1970). *Statistical inference*. Harmondsworth, U.K.: Penguin.
- Spitzer, A.R., Cohen, L.G., Fabrikant, J., & Hallet, M. (1989). A method for determining optimal interelectrode spacing for cerebral topographic mapping. *Electroencephalography and Clinical Neurophysiology*, 72, 355–361.
- St. John, R.C., & Draper, N.R. (1975). D-optimality for regression designs: A review. *Technometrics*, 17, 15–23.
- Tiitinen, H., Sivonen, P., Alku, P., Virtanen, J., & Näätänen, R. (1999). Electromagnetic recordings reveal latency differences in speech and time processing in humans. *Cognitive Brain Research*, 8, 355–363.
- Vaidyanathan, C., & Buckley, K.M. (1997). A sampling theorem for EEG electrode configuration. *IEEE Transactions on Biomedical Engineering*, 44(1), 94–97.
- Vardi, Y., Shepp, L.A., & Kaufman, L. (1985). A statistical model for positron emission tomography. *Journal of the American Statistical Association*, 80, 8–20.
- Waldorp, L.J., Huizenga, H.M., Dolan, C.V., & Molenaar, P.C.M. (2001). Estimated generalized least squares electromagnetic source analysis based on a parametric noise covariance model. *IEEE Transactions on Biomedical Engineering*, 48(6), 737–741.

Manuscript received 16 MAR 1998

Final version received 15 AUG 2001